

Lezione 2

Le molecole di base che
costituiscono la vita

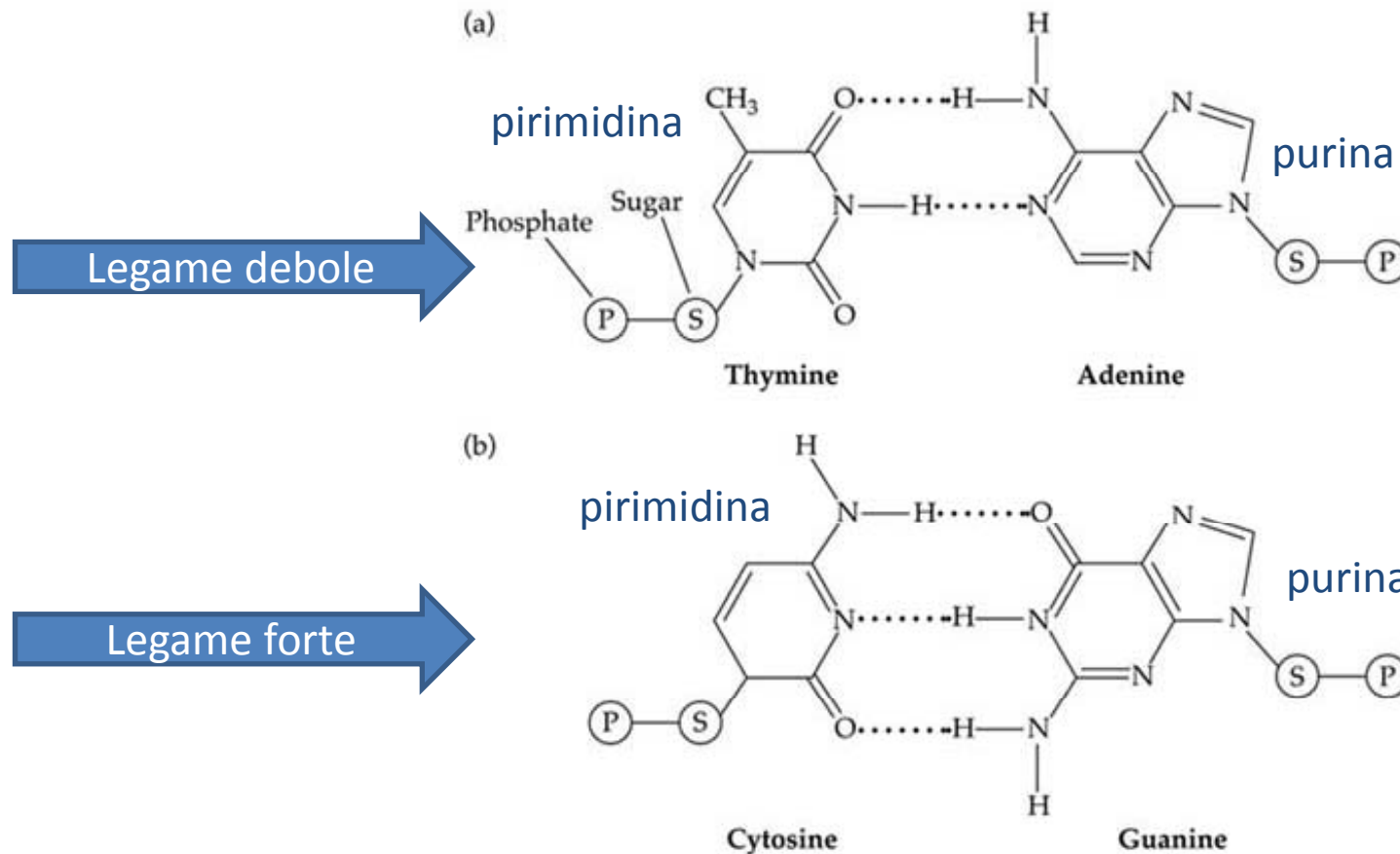
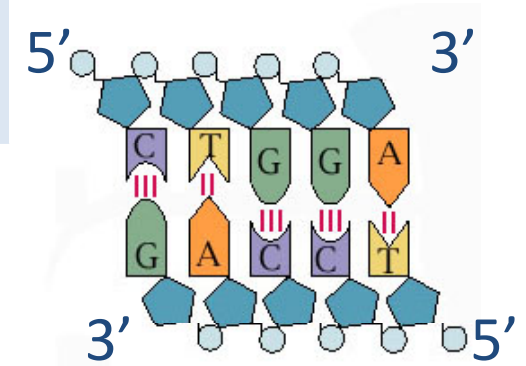
Graur and Li: Capitolo 1

Graur lectures 5-6-7

Geni, codice genetico, mutazioni

Le molecole dell'ereditarietà

L'informazione ereditaria di tutti gli organismi viventi, con l'eccezione di alcuni virus, è a carico della molecola dell'**acido desossiribonucleico (DNA)**.

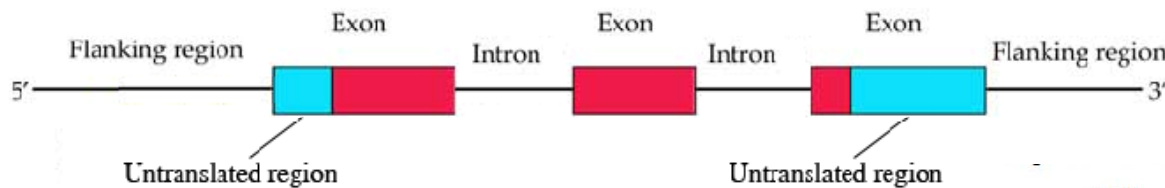


Geni, codice genetico, mutazioni

Definizione “estesa” di gene:

Sequenza di DNA genomico (o di RNA) essenziale per una specifica funzione

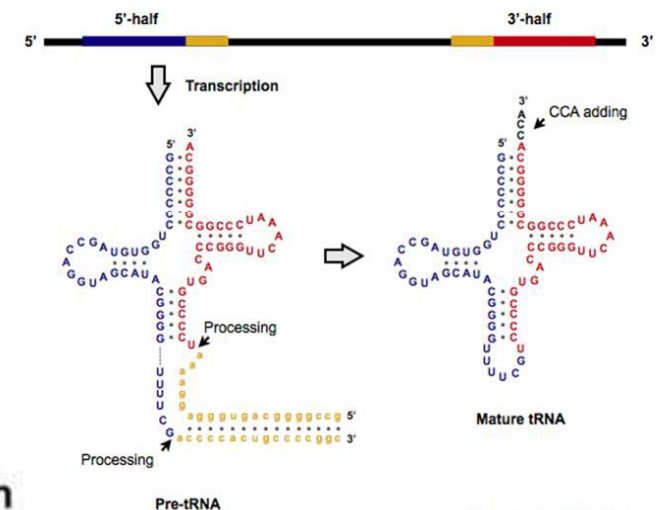
1. Geni che codificano proteine (trascritti > tradotti)



2. Geni che specificano RNA (solo trascritti)

3. Geni non trascritti

4. Pseudogeni (ψ)



Randau et al. 2005. Nature

Geni, codice genetico, mutazioni

Aminoacidi

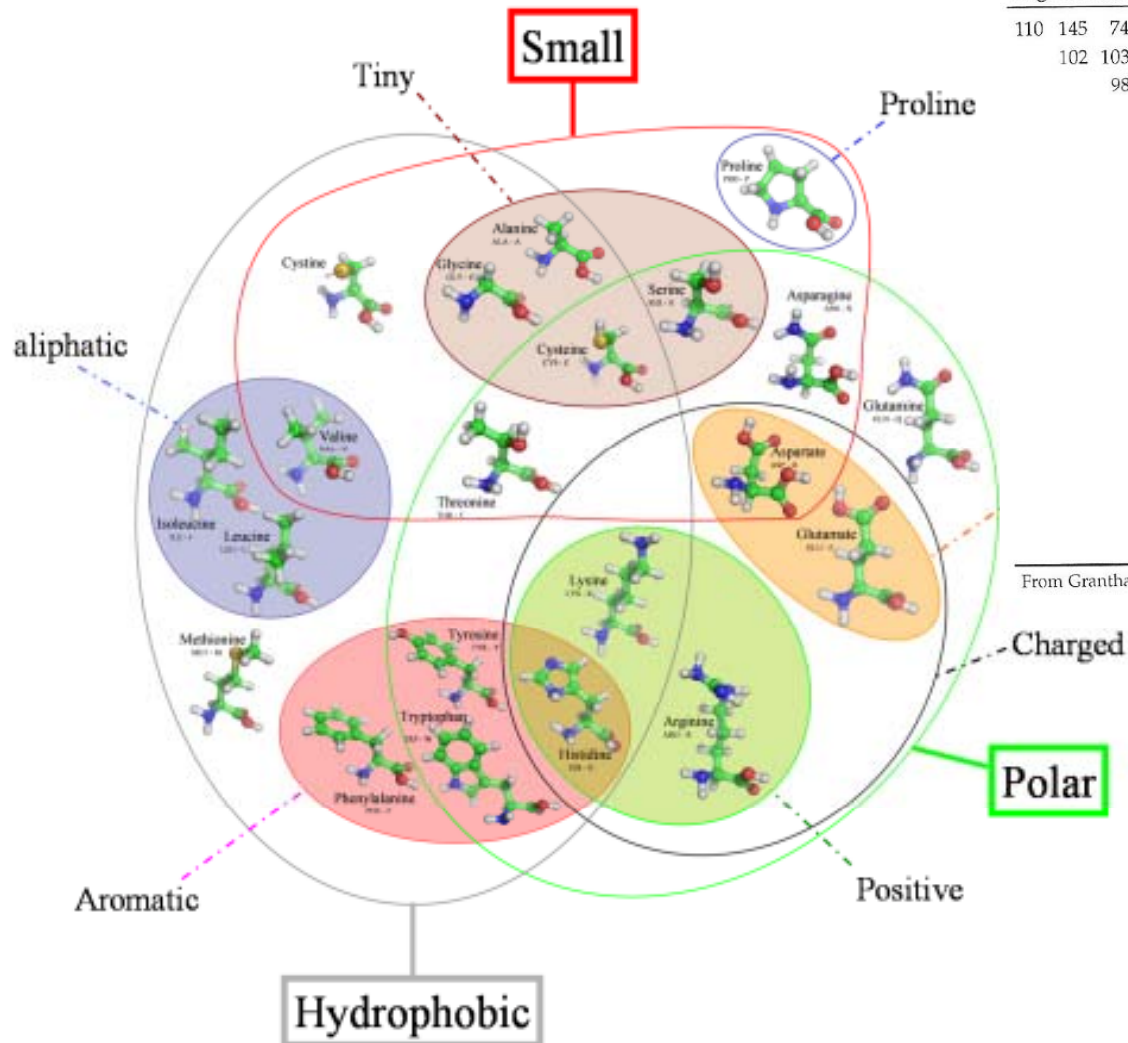


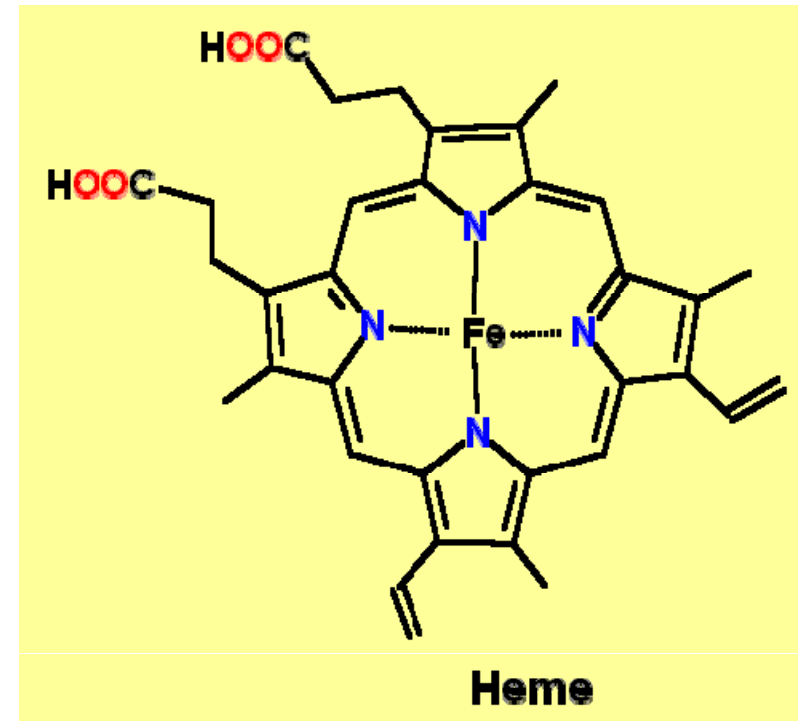
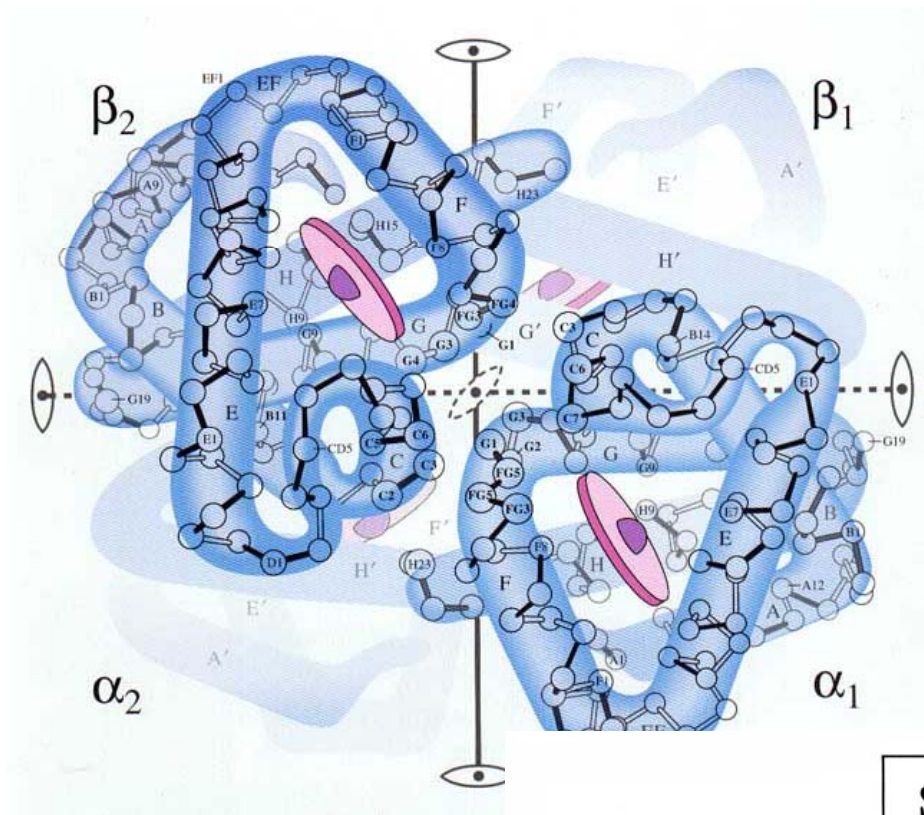
TABLE 4.7 Physicochemical distances between pairs of amino acids^a

Arg	Leu	Pro	Thr	Ala	Val	Gly	Ile	Phe	Tyr	Cys	His	Gln	Asn	Lys	Asp	Glu	Met	Trp	
110	145	74	58	99	124	56	142	155	144	112	89	68	46	121	65	80	135	177	Ser
	102	103	71	112	96	125	97	97	77	180	29	43	86	26	96	54	91	101	Arg
		98	92	96	32	138	5	22	36	198	99	113	153	107	172	138	15	61	Leu
			38	27	68	42	95	114	110	169	77	76	91	103	108	93	87	147	Pro
				58	69	59	89	103	92	149	47	42	65	78	85	63	81	128	Thr
					64	60	94	113	112	195	86	91	111	106	126	107	84	148	Ala
						109	29	50	55	192	84	96	133	97	152	121	21	88	Val
							135	153	147	159	98	87	80	127	94	98	127	184	Gly
								21	33	198	94	109	149	102	168	134	10	61	Ile
									22	205	100	116	158	102	177	140	28	40	Phe
										194	83	99	143	85	160	122	36	37	Tyr
											174	154	139	202	154	170	196	215	Cys
												24	68	32	81	40	87	115	His
													46	53	61	29	101	130	Gln
														94	23	42	142	174	Asn
															101	56	95	110	Lys
																45	160	181	Asp
																	126	152	Glu
																		67	Met

From Grantham (1974).

Geni, codice genetico, mutazioni

Proteine

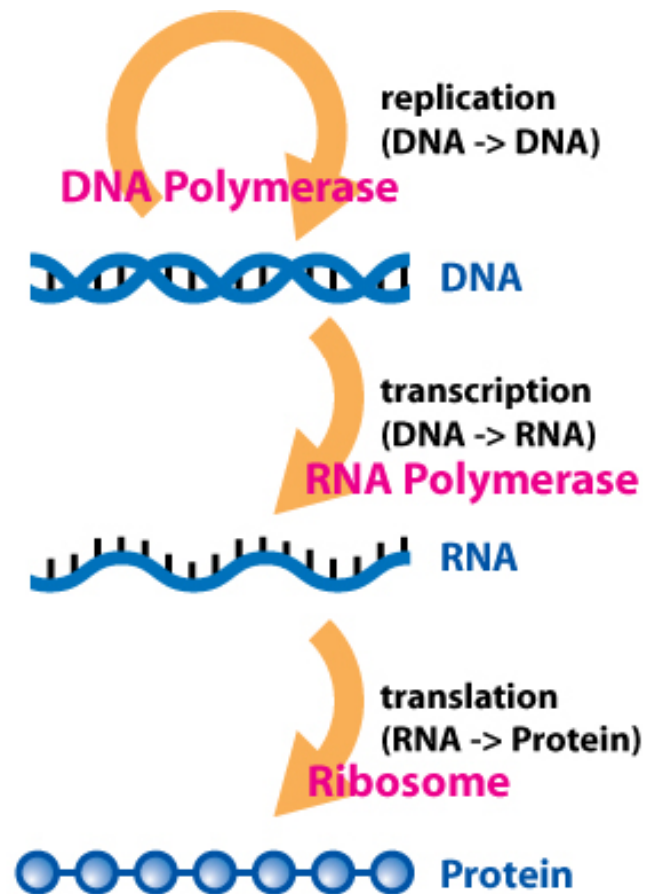


Tasso di
sostituzione per
sito per 10⁹ anni

	surface	heme pocket
α -globin	1.35	0.17
β -globin	2.73	0.24

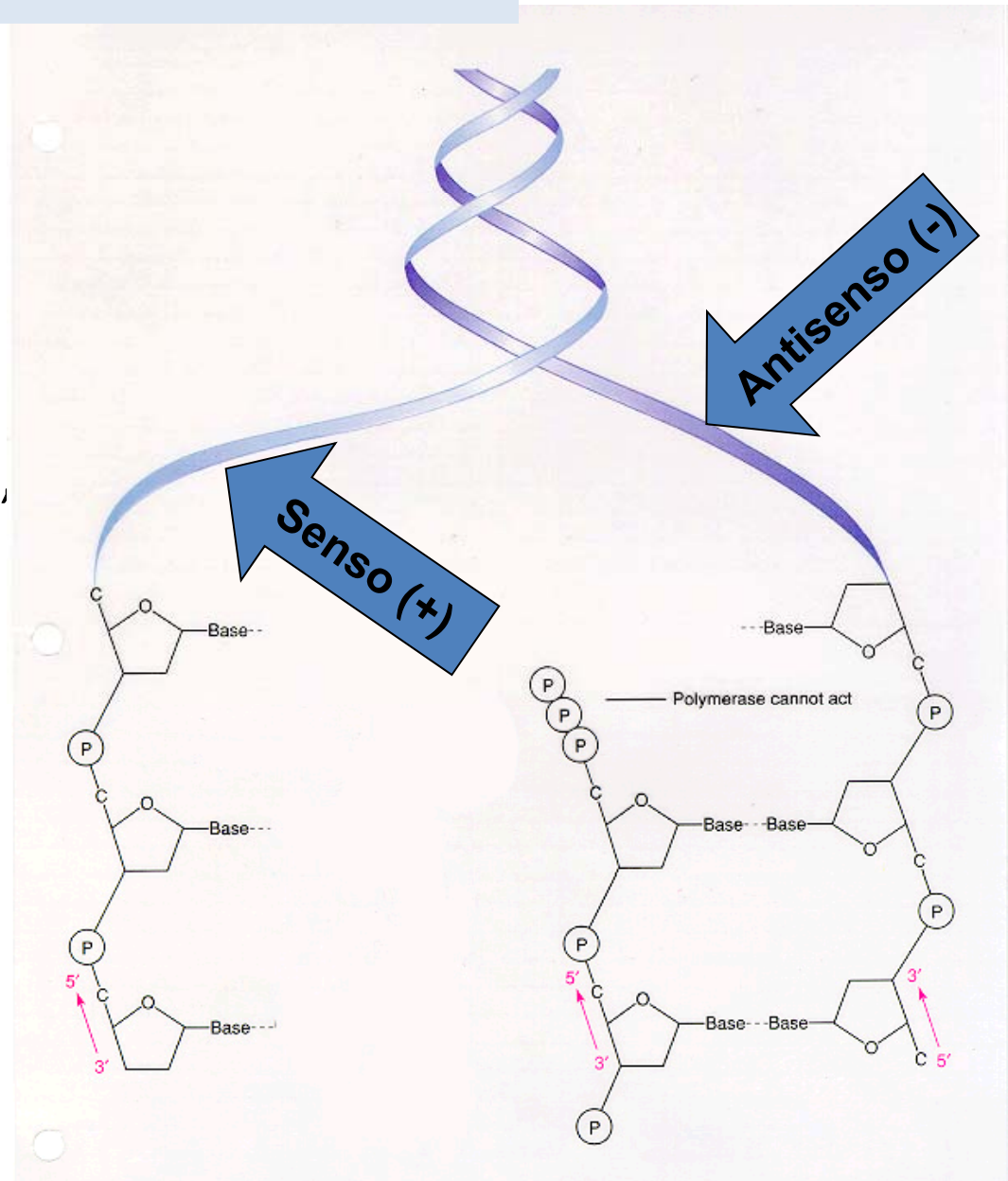
Geni, codice genetico, mutazioni

Il dogma centrale della biologia molecolare: il flusso dell'informazione



Geni, codice genetico, mutazioni

La trascrizione:
il DNA antisenso (strand -) 5' -3'
viene trascritto in un
RNA 3'-5'

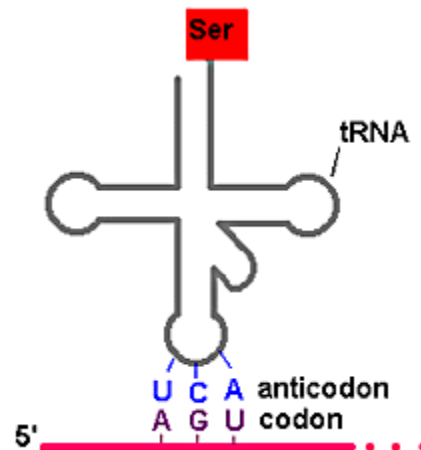


Geni, codice genetico, mutazioni

La trascrizione



La traduzione

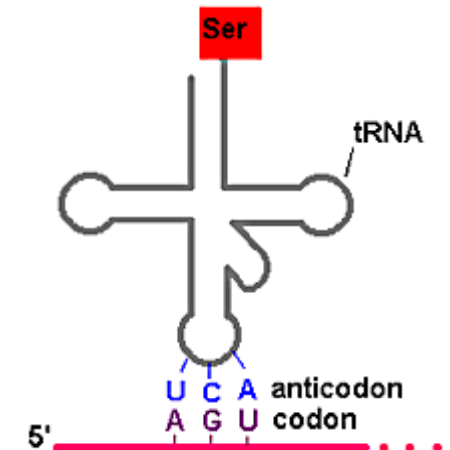


DNA strand +: la stessa tripletta dell'mRNA con T al posto di U

Geni, codice genetico, mutazioni

Il codice genetico universale

		Second Letter					
		U	C	A	G		
1st letter	U	UUU Phe UUC UUA Leu UUG	UCU UCC Ser UCA UCG	UAU Tyr UAC UAA Stop UAG Stop	UGU Cys UGC UGA Stop UGG Trp	U C A G	3rd letter
	C	CUU CUC Leu CUA CUG	CCU CCC Pro CCA CCG	CAU His CAC CAA Gln CAG	CGU CGC Arg CGA CGG	U C A G	
	A	AUU AUC Ile AUA AUG Met	ACU ACC Thr ACA ACG	AAU Asn AAC AAA Lys AAG	AGU Ser AGC AGA Arg AGG	U C A G	
	G	GUU GUC Val GUA GUG	GCU GCC Ala GCA GCG	GAU Asp GAC GAA Glu GAG	GGU GGC Gly GGA GGG	U C A G	



Geni, codice genetico, mutazioni

Il codice genetico organizzato secondo un criterio di degenerazione

Amino Acids with one Codon

AUG UGG

Met Trp

Amino Acids with two codons

AAA AAC CAA CAC GAA GAC UAC UGC UUC
AAG AAU CAG CAU GAG GAU UAU UGU UUU

Lys Asn Gln His Glu Asp Tyr Cys Phe

Amino acid with three codons

AUA
AUC
AUU

Ile

Amino Acids with four codons

ACA ACC CCA CCC GCA GCC GGA GGC GUA
ACG ACG CCG CCG GCG GCG GGG GGG GUG
ACU ACU CCU CCU GCU GCU GGU GGU GUU

Thr Pro Ala Gly Val

Amino Acid with four codons

CGA CGC CGG CGU GUA CUC CUG CUU UCA UCC UCG
AGA AGG UUA UUG UCU UGU AGC AGU

Arg Leu Ser

Geni, codice genetico, mutazioni

Il codice genetico dei mitocondri dei vertebrati

		Second position					
		U	C	A	G		
First position (5'-end)	U	UUU <i>phe</i>	UCU	UAU <i>tyr</i>	UGU <i>cys</i>	U	Third position (3'-end)
		UUC <i>phe</i>	UCC <i>ser</i>	UAC <i>tyr</i>	UGC <i>cys</i>	C	
		UUA <i>leu</i>	UCA <i>ser</i>	UAA <i>Stop</i>	UGA <i>trp</i>	A	
		UUG <i>leu</i>	UCG	UAG <i>Stop</i>	UGG <i>trp</i>	G	
	C	CUU <i>leu</i>	CCU	CAU <i>his</i>	CGU	U	
		CUC <i>leu</i>	CCC <i>pro</i>	CAC <i>his</i>	CGC <i>arg</i>	C	
		CUA <i>leu</i>	CCA <i>pro</i>	CAA <i>gln</i>	CGA <i>arg</i>	A	
		CUG <i>leu</i>	CCG	CAG <i>gln</i>	CGG	G	
	A	AUU <i>ile</i>	ACU	AAU <i>asn</i>	AGU <i>ser</i>	U	
		AUC <i>ile</i>	ACC <i>thr</i>	AAC <i>asn</i>	AGC <i>ser</i>	C	
		AUA <i>met</i>	ACA <i>thr</i>	AAA <i>lys</i>	AGA <i>Stop</i>	A	
		AUG <i>met</i>	ACG	AAG <i>lys</i>	AGG <i>Stop</i>	G	
	G	GUU <i>val</i>	GCU	GAU <i>asp</i>	GGU	U	
		GUC <i>val</i>	GCC <i>ala</i>	GAC <i>asp</i>	GCC <i>gly</i>	C	
		GUA <i>val</i>	GCA <i>ala</i>	GAA <i>glu</i>	GGA <i>gly</i>	A	
		GUG <i>val</i>	GCG	GAG <i>glu</i>	GCG	G	

The Genetic Codes

Compiled by Andrzej (Anjay) Elzanowski and Jim Ostell
National Center for Biotechnology Information (NCBI), Bethesda, Maryland, U.S.A.

Last update of the Genetic Codes: July 07, 2010

The following genetic codes are described here:

- [The Standard Code](#)
- [The Vertebrate Mitochondrial Code](#)
- [The Yeast Mitochondrial Code](#)
- [The Mold, Protozoan, and Coelenterate Mitochondrial Code and the Mycoplasma/Spiroplasma Code](#)
- [The Invertebrate Mitochondrial Code](#)
- [The Ciliate, Dasycladacean and Hexamita Nuclear Code](#)
- [The Echinoderm and Flatworm Mitochondrial Code](#)
- [The Euplotid Nuclear Code](#)
- [The Bacterial, Archaeal and Plant Plastid Code](#)
- [The Alternative Yeast Nuclear Code](#)
- [The Ascidian Mitochondrial Code](#)
- [The Alternative Flatworm Mitochondrial Code](#)
- [Blepharisma Nuclear Code](#)
- [Chlorophycean Mitochondrial Code](#)
- [Trematode Mitochondrial Code](#)
- [Scenedesmus Obliquus Mitochondrial Code](#)
- [Thraustochytrium Mitochondrial Code](#)

<http://www.ncbi.nlm.nih.gov/Taxonomy/Utils/wprintgc.cgi>

Geni, codice genetico, mutazioni

Le mutazioni

Le sequenze di DNA sono solitamente copiate in modo preciso durante la replicazione.

Raramente tuttavia possono avvenire degli errori che originano nuove sequenze. Questi errori si chiamano **mutazioni**.

Geni, codice genetico, mutazioni

Le mutazioni

Da un punto di vista evolutivo una **mutazione** è una sequenza nella linea germinale che differisce dalla sua controparte nelle cellule somatiche, che viene ereditata dalla progenie la quale sarà dunque caratterizzata da una **“novità” genetica**.

Le mutazioni sono quindi la **fonte di variabilità e di novità evolutiva**

Geni, codice genetico, mutazioni

Le mutazioni

(a) AGGCAAACCTACTGGTCTTAT

Transizione
(pur>pur ; pir>pir)

(b) AGGCAA^{*}A^{*}TCTACTGGTCTTAT

Trasversione
(pur>pir ; pir>pur)

sostituzioni

(c) AGGCAAACCTACTG^{*}C^{*}TCTTAT

(d) AGGCAAACCTACTGCAAACAT

ricombinazione

GTCTT

(e) AGGCAA^{ACCTA}CTGGTCTTAT

delezione

(f) AGGCAAACCTACTAAAGCGGTCTTAT

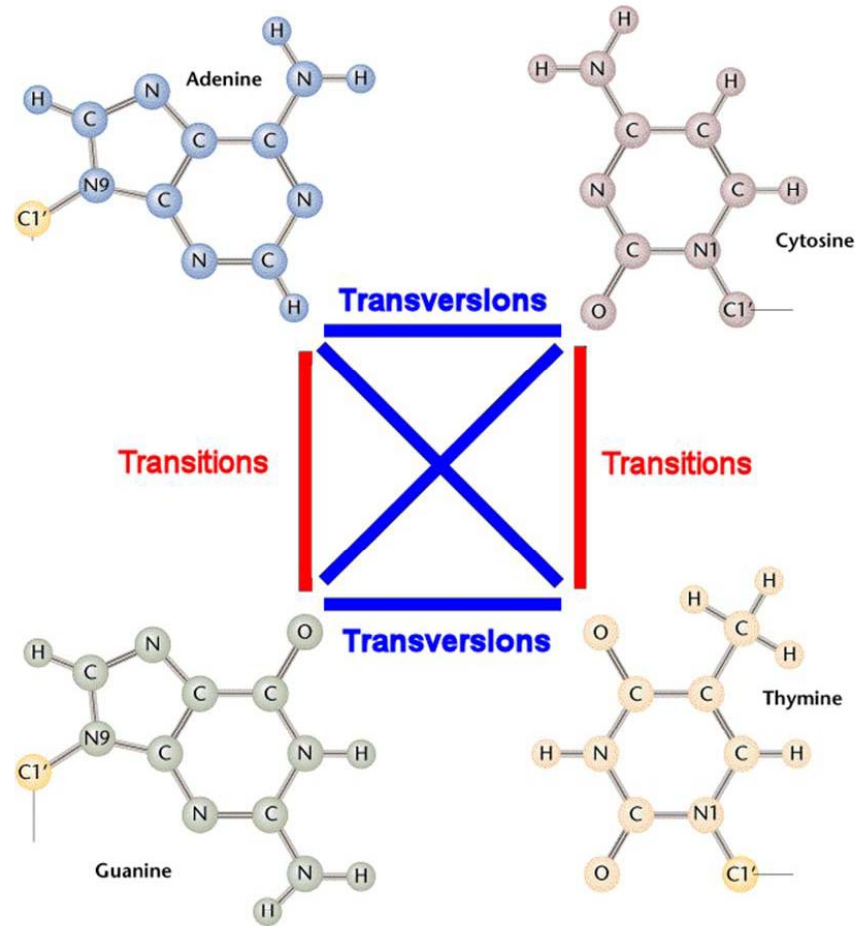
inserzione

(g) AGGTTTGCCTACTGGTCTTAT

inversione

Geni, codice genetico, mutazioni

Le mutazioni



Nucleotide substitutions

Transitions:

$A \rightarrow G$

$G \rightarrow A$

$C \rightarrow T$

$T \rightarrow C$

Transversions:

$A \rightarrow C$

$A \rightarrow T$

$C \rightarrow A$

$C \rightarrow G$

$T \rightarrow A$

$T \rightarrow G$

$G \rightarrow C$

$G \rightarrow T$

Geni, codice genetico, mutazioni

Le mutazioni

(a)	Ile	Cys	Ile	Lys	Ala	Leu	Val	Leu	Leu	Thr
	ATA	TGT	ATA	AAG	GCA	CTG	GTC	CTG	TTA	ACA
							↓	sinonima		
	ATA	TGT	ATA	AAG	GCA	CTG	GTA	CTG	TTA	ACA
	Ile	Cys	Ile	Lys	Ala	Leu	Val	Leu	Leu	Thr
(b)	Ile	Cys	Ile	Lys	Ala	Asn	Val	Leu	Leu	Thr
	ATA	TGT	ATA	AAG	GCA	AAC	GTC	CTG	TTA	ACA
							↓	nonsinonima		
	ATA	TGT	ATA	AAG	GCA	AAC	TTC	CTG	TTA	ACA
	Ile	Cys	Ile	Lys	Ala	Asn	Phe	Leu	Leu	Thr
(c)	Ile	Cys	Ile	Lys	Ala	Asn	Val	Leu	Leu	Thr
	ATA	TGT	ATA	AAG	GCA	AAC	GTC	CTG	TTA	ACA
				↓	nonsense					
	ATA	TGT	ATA	TAG	GCAAACGTCCTGTTAACA					
	Ile	Cys	Ile	Stop						

Geni, codice genetico, mutazioni

Le mutazioni

(a)	Ile	Cys	Ile	Lys	Ala	Leu	Val	Leu	Leu	Thr
	ATA	TGT	ATA	AAG	GCA	CTG	GTC	CTG	TTA	ACA
							↓	sinonima		
	ATA	TGT	ATA	AAG	GCA	CTG	GTA	CTG	TTA	ACA
	Ile	Cys	Ile	Lys	Ala	Leu	Val	Leu	Leu	Thr
(b)	Ile	Cys	Ile	Lys	Ala	Asn	Val	Leu	Leu	Thr
	ATA	TGT	ATA	AAG	GCA	AAC	GTC	CTG	TTA	ACA
							↓	nonsinonima		
	ATA	TGT	ATA	AAG	GCA	AAC	TTC	CTG	TTA	ACA
	Ile	Cys	Ile	Lys	Ala	Asn	Phe	Leu	Leu	Thr
(c)	Ile	Cys	Ile	Lys	Ala	Asn	Val	Leu	Leu	Thr
	ATA	TGT	ATA	AAG	GCA	AAC	GTC	CTG	TTA	ACA
				↓	nonsense					
	ATA	TGT	ATA	TAG	GCAAACGTCCTGTTAACA					
	Ile	Cys	Ile	Stop						

Geni, codice genetico, mutazioni

Le mutazioni

Ogni codone codificante un AA può mutare in altri 9 codoni attraverso sostituzioni di un singolo nucleotide.

Esempio:

CCU (Pro)

6 possibili sostituzioni nonsinonime

UCU (Ser)

ACU (Thr)

GCU (Ala)

CUU (Leu)

CAU (His)

CGU (Arg)

3 possibili sostituzioni sinonime

CCC

CCA

CCG

Geni, codice genetico, mutazioni

Le mutazioni

Ogni codone codificante un AA può mutare in altri 9 codoni attraverso sostituzioni di un singolo nucleotide.

		Second Letter				
		U	C	A	G	
1st letter	U	UUU Phe UUC UUA Leu UUG	UCU UCC Ser UCA UCG	UAU Tyr UAC UAA Stop UAG Stop	UGU Cys UGC UGA Stop UGG Trp	U C A G
	C	CUU CUC Leu CUA CUG	CCU CCC Pro CCA CCG	CAU His CAC CAA Gln CAG	CGU CGC Arg CGA CGG	U C A G
	A	AUU AUC Ile AUA AUG Met	ACU ACC Thr ACA ACG	AAU Asn AAC AAA Lys AAG	AGU Ser AGC AGA Arg AGG	U C A G
	G	GUU GUC Val GUA GUG	GCU GCC Ala GCA GCG	GAU Asp GAC GAA Glu GAG	GGU GGC Gly GGA GGG	U C A G

61 codoni

“senso”



61x9 = 549

possibili

sostituzioni

nucleotidiche

Geni, codice genetico, mutazioni

Le mutazioni

Se assumiamo che

1. Tutti i codoni siano ugualmente presenti nelle regioni codificanti
2. Ogni sito abbia la stessa probabilità di mutare

In un gene codificante qualunque ci aspettiamo una frequenza relativa dei diversi tipi di sostituzioni come in tabella

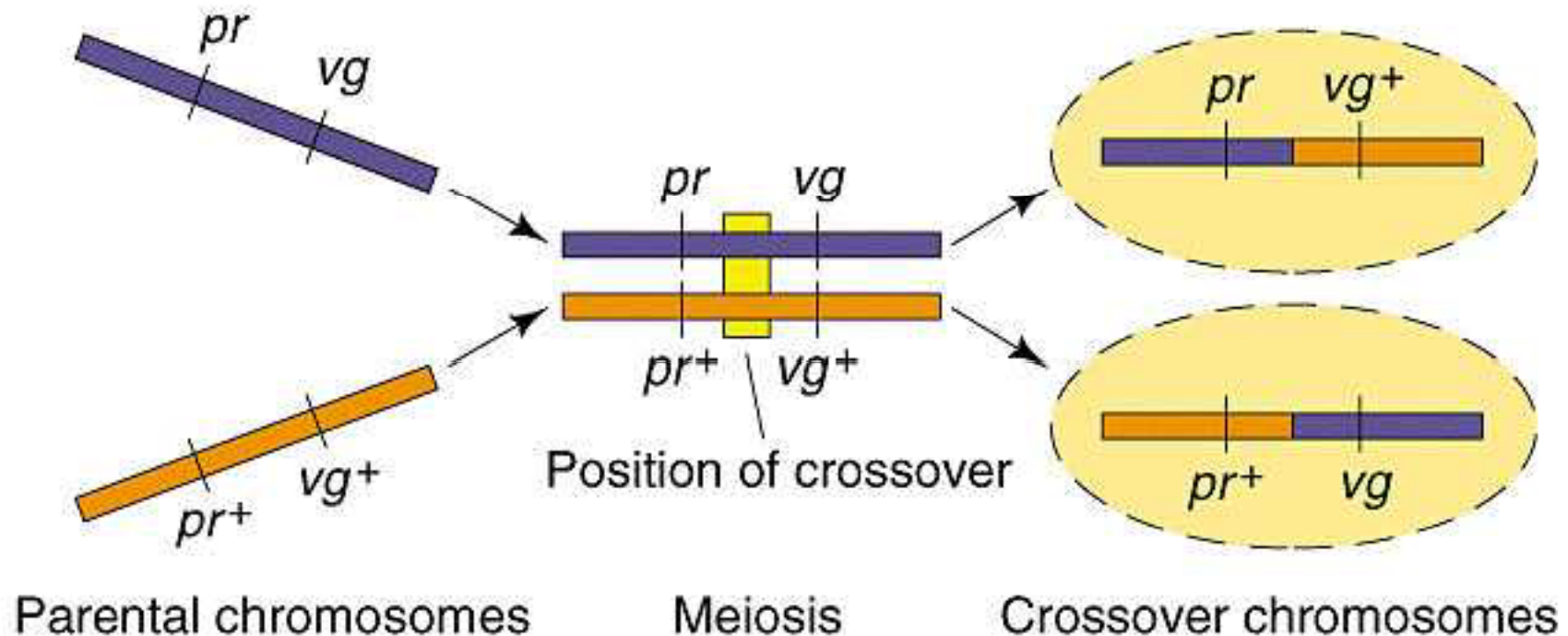
Alcune caratteristiche importanti:

- Circa il 70% dei cambiamenti in 3° base sono sinonimi
- Il 100% dei cambiamenti in 2° base sono nonsinonimi
- Il 96% dei cambiamenti in 1° base sono nonsinonimi

sostituzioni	numero	Frequenza
Totali (1,2,3 base)	549	100
Sinonime	134	25
Nonsinonime	415	75
Missenso (non senso)	392 (23)	71 (4)
Totali (1 base)	183	100
Sinonime	8	4
Nonsinonime	175	96
Missenso (non senso)	166 (9)	91 (5)
Totali (2 base)	183	100
Sinonime	0	0
Nonsinonime	183	100
Missenso (non senso)	176 (7)	96 (4)
Totali (3 base)	183	100
Sinonime	126	69
Nonsinonime	57	31
Missenso (non senso)	50 (7)	27 (4)

Geni, codice genetico, mutazioni

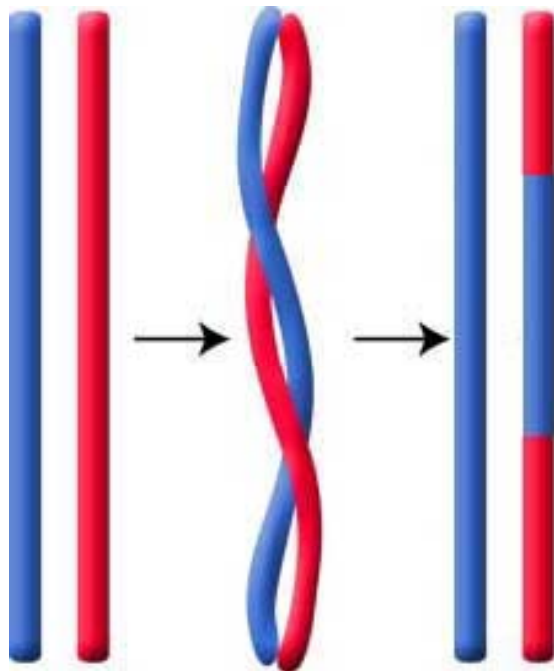
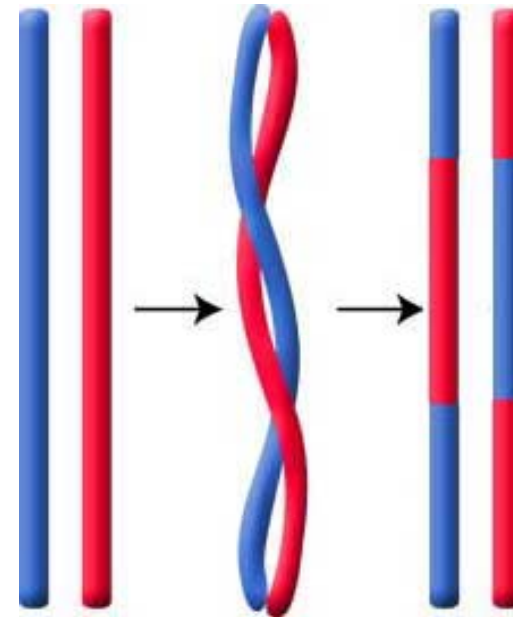
Altre fonti di variabilità: la ricombinazione



Geni, codice genetico, mutazioni

Recombination

1. Crossing-over or Reciprocal recombination.



2. Gene conversion or nonreciprocal recombination.

Geni, codice genetico, mutazioni

La ricombinazione reciproca è un potente mezzo di generazione della variabilità

5'—AACT—3' and 5'—CTTG—3' → **6 possibili nuove** sequenze:

5'—ATTG—3'

5'—CACT—3'

5'—AATG—3'

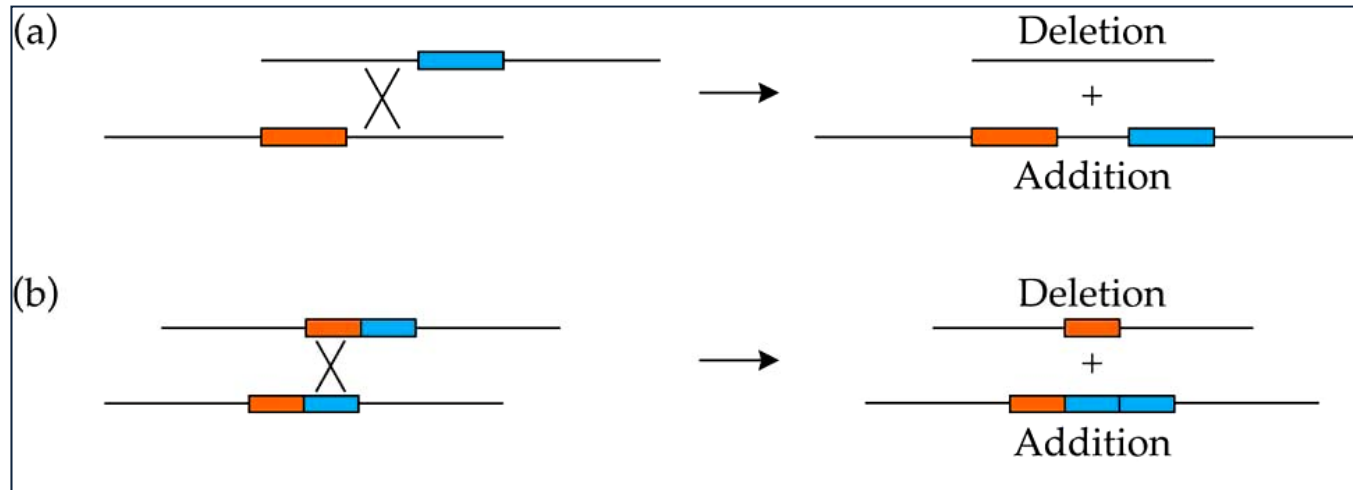
5'—CTCT—3'

5'—AACG—3'

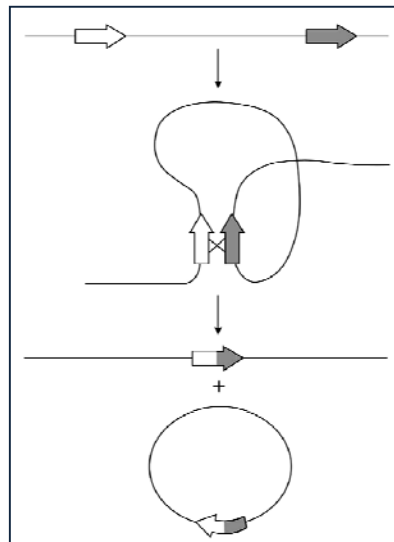
5'—CTTT—3'

Geni, codice genetico, mutazioni

Inserzioni e delezioni



Crossing over
ineguale



Delezione
“intra strand”

(a) Lys Ala Leu Val Leu Leu Thr Ile Cys Ile Stop
 AAG GCA CTG GTC CTG TTA ACA ATA TGT ATA TAA TACCATCGCAATATGAAAATC
 ↓
 G

Terminazione prematura per delezione

AAG GCA CTG TCC TGT TAA CAATATGTATATAATACCATCGCAATATGAAAATC
 Lys Ala Leu Phe Cys Stop

(b) Lys Ala Leu Val Leu Leu Thr Ile Cys Ile Stop
 AAG GCA CTG GTC CTG TTA ACA ATA TGT ATA TAA TACCATCGCAATATGAAAATC
 ↓
 A

Perdita di un codone di stop per delezione

AAG GCA CTG GTC CTG TTA ACA ATA TGT ATT AAT ACC ATC GCA ATA TGA AAA
 Lys Ala Leu Val Leu Leu Thr Ile Cys Ile Asn Thr Ile Ala Ile Stop

(c) Lys Ala Axn Val Leu Leu Thr Ile Cys Ile Stop
 AAG GCA AAC GTC CTG TTA ACA ATA TGT ATA TAA TACCATCGCAATAGGG
 ↑
 G

Perdita di un codone di stop per inserzione

AAG GCA AAC GGT CCT GTT AAC AAT ATG TAT ATA ATA CCA TCG CAA TAG GG
 Lys Ala Asn Gly Pro Val Asn Asn Met Tyr Ile Ile Pro Ser Gln Stop

(d) Lys Ala Asn Val Leu Leu Thr Ile Cys Ile Stop
 AAG GCA AAC GTC CTG TTA ACA ATA TGT ATA TAA TACCATCGCAATAGGG
 ↑
 GA

Terminazione prematura per inserzione

AAG GCA AAC GAG TCC TGT TAA CAATATGTATATAATACCATCGCAATAGGG
 Lys Ala Asn Glu Ser Cys Stop

F
r
a
m
e
s
s
h
i
f
t

Geni, codice genetico, mutazioni

Inserzioni e delezioni

Nel confronto tra due sequenze è impossibile capire se ci sia stata una delezione in una delle due o una inserzione nell'altra

INserzioni de **DEL**ezioni vengono in generale chiamate **INDELS**